# 大数据时代哲学社会科学学术成果评价：问题、策略及指标体系*

李品 杨建林
南京大学信息管理学院 南京 210023 江苏省数据工程与知识服务重点实验室 南京 210023

**摘要：**［目的/意义］学术成果评价是推进哲学社会科学健康发展的重要因素，长期以来，传统同行评议和科学计量学方法在哲学社会科学学术成果评价中暴露出的问题越来越多，挑战固有方法的统治地位、开发有针对性的评价策略已刻不容缓。文章试图探索大数据环境下哲社学术成果评价的变革及其实现，特别是提出了大数据思维下的哲社学术成果的评价指标体系。［方法/过程］基于比较分析和综合分析，对传统哲学社会科学评价方法的弊端进行分析，然后对大数据给哲学社会科学评价带来的改变进行分析，最后提出基于大数据环境的哲学社会科学评价策略和指标体系。［结果/结论］提出大数据时代哲学社会科学学术成果评价的策略：由引文著录分析转向多维度的引用内容与行为分析，由面向成果的阶段性静态评价转向以"学术活动"为中心的全过程动态评价，由学术影响力评价转向学术价值和社会效益评价。在此基础上，构建由两个一级指标、5个二级指标和34个三级指标组成的大数据背景下哲学社会科学学术成果评价指标体系。

**关键词：**哲学社会科学 学术成果 成果评价 大数据
**分类号：**G251
**DOI**：10.13266/j.issn.0252-3116.2018.16.001



## 1 引言

长期以来，学术评价一直处于定量评价与定性评价的分裂中。定量评价因其客观有限性、公正性不足、专业缺位、结论简单等弊端饱受诟病，南京大学苏新宁教授指出，极端量化、唯量化的学术评价应当终止[1]；大连理工大学李冲等也指出，量化评价因学科的不同而普遍适用性有限[2]；而定性评价也因其主观随意性、低效性、评价过程隐蔽性、结果难以复证和监督的缺陷等饱受质疑，虽有定性与定量评价相结合的方法作为互补和填充，但因实施同行评议的学术共同体与第三方评价机构这两个定性与定量评价主体难以融合、各自为政，结合的效果始终差强人意。当下的问题是，第三方评价机构由于其缺失专业性而无法参与定性评价，即便是专注于定量评价，也因其无法做出专业性结论这一短板，使其评价有失完整。但是学术共同体却未必不能在定量评价上下功夫。当前学术界公认哲学社会科学学术成果评价是学术共同体的定性评价而非第三方机构的定量评价，但这是前大数据时代的共识。

大数据时代，这些问题当有所改变。大数据所带来的"基于云计算的数据处理与应用模式，通过数据的集成共享、交叉复用形成的智力资源和知识服务能力"[3]不仅可以拓展定量分析的数据类型和研究范畴，还可以借助数据挖掘与分析等前期辅助作用，使学术共同体的定性评价突破人力与时间的界限，使海量的定性分析和高效的价值判断成为可能，同时，还可以监督其主观性，佐证其公信力。换言之，大数据时代，使定量评价以更为专业的姿态融入到哲学社会科学的学术成果评价成为可能，也使定性评价的高效性和公信力得以保障。定性与定量评价深入的无缝对接与融合在大数据时代将成为新的评价方式。基于此，国际上量化评价方法已呈现出了新的发展趋势：深度质量化和语义化趋向、替代计量数据快速兴起[4]。可见，大数据之于学术评价的渗透已渐渐风生水起。

习近平总书记在哲学社会科学工作座谈会上针对

* 本文系江苏省社会科学基金重大项目"习近平总书记构建中国特色哲学社会科学重大命题研究"（项目编号：16ZD004）研究成果之一。
作者简介：李品（ORCID：0000-0002-0762-177X），博士研究生，E-mail：376341195@qq.com；杨建林（ORCID：0000-0001-6300-1164），教授，博士，博士生导师。
收稿日期：2017-12-13 修回日期：2018-04-23 本文起止页码：5-14 本文责任编辑：易飞

我国哲学社会科学领域存在的问题指出,我国哲学社会科学的“学术评价体系不够科学”。中国社会科学杂志社研究室学者彭秋归认为,学术性和意识形态属性是哲学社会科学研究两个最基本的属性,学术标准和政治标准是哲社评价中两个最基本的要求[5]。哲学社会科学(以下简称“哲社”)学术成果的特殊性,是其在评价研究中被单独剥离出来讨论的主要原因。学界对哲学(人文)社会科学评价问题也给予了一定重视,如大连理工大学姜春林认为,人文社会科学研究成果评价要基于成果的内在属性而展开[6];重庆邮电大学代金平提出,哲学社会科学评价除注重学术创新性和学术影响力外,更要重视研究者的基本价值取向,注重成果的价值导向[7]。

传统的同行评议和科学计量学方法在哲社学术成果评价中已显得力不从心。大数据时代为现有方法的突破提供了新的思维模式(基于多元、分布式数据的关联)以及推动这种思维模式得以实现的数据环境(提供不同格式、不同学术活动的学术成果发布平台)与技术支撑(文本挖掘技术、语义关联技术等)。基于此,本文提出了大数据思维下的哲社学术成果评价策略,重点探索大数据环境下哲社学术成果评价的变革及其实现,特别是提出了大数据思维下的哲社学术成果的评价指标体系。以此使哲社成果评价从注重以影响力等外围指标为主深入、拓展到引用内容、社会价值等深层次的内容指标,从基于期刊论文、专著等数据的评价拓展到各种类型和表现形式的学术活动评价,从而更加深刻而全面地揭示学者学术表现。



## 2 哲社学术成果评价存在的问题与思考

当前,我国各学科学术成果评价主要由以质性评估为核心的同行评议法和以量化评估为手段的科学计量法两种方法构成[8]。日本学者 K. Selvarajoo 指出,应用科学计量学来评价学术成就正在误导各种年龄阶段的学者[9],而同行评议法须建立在专家的价值观和思维方式之上,受主观因素影响较大[10],这两种方法本身的缺陷在国内外已被较普遍地认识。

不同于自然科学(以下简称自科)注重探索自然界的真理,可以在世界范围内采用共通的标准,哲社研究对象很大程度是人和社会,这就决定了哲社学术成果的特殊性:①成果内容以思想性、意识形态性、价值导向性为主要特征。先不考虑专家的社会文化影响,本身这些内容评价就对同行评议专家的学术素养和思维认知提出了很高的要求,而科学计量学指标更是几乎无能为力;②成果的传播与应用具有明显的影响延时性、成果转化模糊性、作用价值体验性和非正式学术交流突出性等特征。多来自于“学术圈内”的理论派同行专家显然只凭自身直觉和经验无法真正深刻而又客观地分析成果中这些问题的表现,而以引文、发文,甚至基于替代计量学(altmetrics)中下载、转载等各项仅流于表面的计量指标,更难以对此做出有效评判;③哲社学者受感性思维和社会环境变化影响显著,这就要求不仅要对其阶段性成果存量进行分析,还须建立长期的面向整个研究过程和学术生涯的评价机制,这是现有同行评议机制和科学计量学评价方法无法企及的。

综上,现有同行评议和科学计量学方法除本身所固有的缺陷影响着哲社学术成果评价的科学性和客观性外,这两种方法也因管理与技术上无法克服的障碍,而没有针对哲社成果的特殊性采取有效的措施来应对,由此更加强化了当前对哲社学术成果认知局限和结论偏差的风险。具体表现在以下方面:

### 2.1 简单线性关联的引文著录分析忽略了引用的复杂含义

引文分析作为学术成果评价的重要方法已出现 50 余年,科学引文索引数据库的出现为其提供了强大的应用动力。其实,基于著录信息的引文分析从产生开始就受到质疑,1986 年,T. A. Brooks 认为,引文分析法在没有搞清引文动机、引文行为之前就开始被使用了[11]。引文著录分析通常是用来描述文献或作者之间的相互影响,这种影响没有任何的指向性和强度性考量,过程中还隐含一个假设:每篇被引文献对施引文献均具有相同的价值[12],显然这是不合理的。实际上,传统引文著录分析主要想回答两个问题:是否两篇文章具有相互关联性? 文章被引用次数是多少? 这就将学术影响力与引用次数等同起来。H. Small 曾指出,没有任何证据表明高被引论文具有高影响力[13]。引文分析存在缺陷的本质原因是将文献间相互引用的关系理解为简单的线性关系,而忽略了引用的下列复杂含义:

#### 2.1.1 引文行为的复杂性赋予了引用率丰富的内涵

引文行为有其范围性,其上限(即对实际参考文献准确且充分引用)和下限之间可能出现不相关或弱相关引文[14],其中引用动机(如非正常的动机[15])、引用内容(如方法类论文占优[16])和引用类型(正引、反引、明引、暗引等)是导致这一现象的关键要素。多年来虽然不少成果对引文行为进行了一定尝试性分析,如 E. Garfield 提出 15 种引用原因,指出被引用次数无法揭

示出作者的引用动机[17]。但因样本量选取（大小、结构）、分析方法和文献收集不完整等难以克服的技术性缺陷，再加之引文行为本身的社会文化复杂性（心理层面、认知层面和动机层面等）问题，始终没能够提出切实可行的有效解决方案。

2.1.2 引用率数据模糊了不同学科间的差异 如理科和工科、人文学科与社会科学、应用学科与基础学科存在显著差异[18]。然而，因学科之间关系的复杂性、学科内部中研究问题的不同（如基础性研究、应用性研究等），导致只关注数据的引文著录分析无法区别学科间的引用差异。尤其是在哲社领域，很多学科的论文无论有多优秀其引用率都有可能不高，这是由学科本身的特征决定的。而科学共同体更多地关注高被引文献，忽视了低被引或零被引文献[19]。

2.1.3 引用过程的时滞性降低了引文分析的效率 W. Glanzel 等研究结果显示：整体上超过 80% 的文献是发表 3 年内首次被引，超过 90% 的文献是发表后 5 年内首次被引[20]；与此类似，荷兰科学计量学家 A. F. J. Van Raan 发现了"睡美人"现象[21]，均有力地证明了文献被引用的时滞性问题，这种时滞会对科研成果影响力、科研人员评价造成拖延，以至于无法真正反映当时当下的真实科研情况，尤其是对于激励年轻科研人员也十分不利。而科学研究中非正式交流的兴起（如学术博客、预印本、学术社区等），使得大量资料在正式发表前已通过非正式渠道广泛传播[22]，非正式交流已成为科研成果传播和影响力深化的重要方式之一，当前仅局限于以文献为基础的引文分析会因在此方面的缺陷，使分析效率大大降低。

### 2.2 面向成果的阶段性静态评价存在以偏概全的风险

在非正式学术交流发展迅猛的背景下，哲社学术成果不仅体现在以文献为载体的表达形式中，而且广泛地存在于各种非正式交流语境中，非正式学术交流具有显著的时效性和动态性，需要实时跟踪分析，静态固化形态的成果分析，显然存在片面性。不仅如此，学术活动和学术生涯等这种以静态成果形式无法表达的动态进程性工作，对于哲社这种"慢热型"和"沉淀型"的特点至关重要，对这些要素进行全面分析，是发现人才、培养人才，进而产生更多优秀学术成果的重要保障。而当前局限于某一阶段并只关注具体成果的评价方式，显然无法实现这一目标。

### 2.3 学术影响力评价无法客观洞察学术成果的学术价值和社会效益

知识经济时代、创新驱动发展战略下，科学研究对于社会发展的重要作用被寄予厚望。哲社研究不仅要产出高影响力的学术成果，更重要的是这些成果能够体现学术价值和产生社会效益。学术价值体现在学术的创新性、前沿性和对学术研究的推动作用，不同于学术影响力大多局限于表面的学术传播与传承，学术价值更多体现的是思想、理念与内容上的深度引导，是加速科学发展的关键；社会效益的核心则是学术成果在推动社会实践（如政策制定、社会管理与发展等）发展中的指导作用。

传统科学计量学指标局限在"学术圈内"的表面影响力评价显然不适应时代发展要求，正如 M. Taylor 指出，当前的科学计量学指标对于理解学术成果在社会中的影响十分有限[23]。尤其是社会发展的复杂性日益突出，全球变暖、食品安全、生态服务缺失、社会发展与管理中暴露的种种矛盾等均需要科学研究对此做出重要贡献，停留于出版物及其引用等指标分析的学术成果评价，很难真实反映学术成果推动社会发展的贡献，正如 J. Spaapen 等所提出的，学术成果评价要超越科学绩效指标[24]。同行评议和学术共同体评价虽然一定程度上改善了科学计量学指标的缺陷，但其不易操控的主观性以及网络环境下资源分布的广泛性导致资源获取与认知的局限，同样无法真正科学、客观、全面地揭示学术成果的学术价值和社会效益。

## 3 基于大数据思维的哲社学术成果评价策略

大数据时代，以各种载体、各种形态存在的各类动、静态学术活动数据可以被深度聚合，并进行语义层面的关联分析。将大数据思维、技术与方法应用于哲社学术成果评价中，可以有效克服当前相关评价的局限。那么，大数据时代我们应该如何建立与之相匹配的哲社学术成果评价策略？

### 3.1 由引文著录分析转向多维度的引用内容与行为分析

相对于自科学术成果内容的确定性和明确性，哲社学术成果语境、语义要复杂得多，引文著录分析这种流于形式和表面的评价方法，无法反映出哲社思想性、意识形态性和价值性内容；不仅如此，哲社学者更容易脱离文献而通过非正式交流手段进行思想碰撞。大数据使面向引用内容的语义分析和融合多种媒介的非正式交流分析成为可能。

3.1.1 基于全文文本的引用内容与行为分析全面深刻地揭示了引用的复杂意义 在引用内容分析界定

上,H. Small 最具影响力,他用“citation context”来表述引用内容,意指引用位置周围的文本内容[25],强调了引用位置、引用语境的重要性。研究性论文通常具有清晰的逻辑结构,各模块实际上在论文中均具有不同级别的重要性,前言(引言)、方法、过程、结论等各部分的重要性级别,决定了引用发生在相应位置中被引文献的重要性大小;而引用语境通过引用位置前后词汇诸如“基于”“在…基础上”(功能大)以及“拓展”(功能重要)、“比较”(功能一般)、“历史背景”(功能不重要)等相关语词的表达[26],显示了引用功能大小。引用内容分析实际上是通过引文功能、引文重要性、情感倾向和引用动机 4 个方面揭示引用意图和被评价对象的影响力,从而更加深刻地洞悉和挖掘引用的复杂性所潜藏的价值。

引用内容分析建立在两个前提之上:一是在不同语境下引用被赋予不同的权重,这主要通过句法和语义分析来实现;二是定性与定量相互融合[27],即内容分析法和计量分析法相互补充。权重大小反映了被引用内容的重要性高低以及功能大小;定性与定量方法的相互补充不仅可以通过语义特征描述引用极性和动机,还可以对它们进行量化累积角度的重要性考察。因此,引用内容分析通过话语分析(语境、语义),从语言学角度揭示引用内容的情感、认知、态度及其重要性和功能,从而可以客观揭示作者的引用意图,展示引用对于自身研究工作的功能,并且能够更加全面深刻地揭示科学技术发展特征(探测研究内容的突破、转移和变革[28]),更加客观真实地考察被评价对象的影响力。



#### 3.1.2 大数据思维下引用内容与行为分析的实现

大数据的兴起,为基于全文信息的引用内容分析提供了基本技术和平台保障,为语义分析和非正式交流科研影响力分析提供了可能。尤其是随着文献数据库建设的日趋完善和信息与数据处理技术的深入发展,越来越多可供全文解析的数据库得以陆续开发,例如 arXiv、Citeseer 等,包括国内的 CNKI 也出版了 Html 格式的全文文本,这些为文献结构与内容的全文解析和文本挖掘提供了十分便利的条件,从而为深入全文的引文内容分析提供了重要的数据基础。同时,社会计算能力的进步和数字图书馆、机构知识库的发展,为识别跨领域的特定引用模式提供了依据,从而为基于自然语言处理的引用内容与行为分析提供了可行的解决方案;此外,在此环境下,可以开发可扩展文本挖掘算法,来更加深入和全面地抽取大范围文献集合中隐藏的语义关联。

(1)通过语义分析实现引用功能探测。语义分析通常可包括两个层面:一是语义贡献分析,主要是指被引文献在施引文献中的贡献性,如重要或不太重要的贡献、肯定或否定型贡献等[29]。这需要对引用位置及其周围语境、代表功能性内容的特定语词进行挖掘分析。二是通过自然语言处理和大数据技术,分析引用动机、情感等隐含性引用行为,挖掘科研信息中存在的重要观点。完成引用内容的语义分析,需要建立引用功能、重要性和动机、情感等相关分类目录及其权重级别,并建立相对应的语料库。语义分析主要步骤包括:建立分析目标,收集、组织数据,明确分析要素构成及特征并依此建立分析类目,对内容语义进行编码处理以及数据统计与检验、解释说明,等等[30]。语义分析可通过机器学习、自然语言处理、文本挖掘、情感分析和可视化分析等相关技术加以实现。

(2)基于非正式交流的科研影响力分析。非正式交流因其克服了传统文献系统交流的时滞性缺陷,使科研交流的实效性增强,而被国内外学者广泛认可。如,B. Cronin 指出,非正式交流是科学发展的源泉[31];J. Price 认为科学家约有 80% 的资料通过非正式交流这种“无形学院”获得[32];黄水清等也充分肯定了非正式交流的价值[33]。2010 年,Altmetrics 的诞生推动了非正式交流在科研影响力评价中的实际应用,是典型的大数据时代下评价科研影响力的产物。

美国科学学与创新政策项目主任 L. Julia 认为,为更好地进行科学评价,应该收集科学家科学研究工作全过程、全范围的数据[34]。即是说,科研活动全过程、全方位数据均应成为学术成果评价的重要组成部分。Web2.0/3.0 环境下,非正式交流的形式日益丰富,OA 期刊、OA 仓储、学术博客、网络社区、即时交流工具等蓬勃兴起,这些载体不仅存储了大量科研内容信息,同时也留下了丰富的科研行为踪迹数据,对这些载体中所呈现的学术内容、学者科研行为轨迹等各种类型和结构的数据进行时空维度的集成与挖掘,对这种科研交流模式下的学者相互影响与关联进行识别,是大数据应用于科研影响力评价的重要方式。

### 3.2 由面向成果的阶段性静态评价转向以“学术活动”为中心的全过程动态评价

相比自科学者科研过程的循序渐近性、浓厚的理性思维和表现形式的固定性,哲社学者受社会环境影响显著,科研行为很多时候因社会环境的变化而发生循环往复,感性思维和经验性在此过程中会起到重要作用,成果表现形式多样化。简单地以传统形式的产

李品，杨建林. 大数据时代哲学社会科学学术成果评价：问题、策略及指标体系[J]. 图书情报工作,2018,62(16):5－14.

出成果为主导来评价哲社学者显然是存在偏差的。大数据环境下，使面向“学术活动”的科研活动过程评价成为可能，将科研工作者的学术活动以及学术生涯过程进行全方位关联分析，不仅能够客观评价他们当前的科研能力，而且可以有效地发现他们未来的科研发展空间，助其自我认知，激发科研斗志。

3.2.1　学术活动大数据的关联分析　学术活动是学者开展与学术相关的研究、交流与传播活动的总称。随着科研工作的开放性发展，学术活动的频率在不断提高，形式也日益丰富。在此过程中会产生丰富且具有针对性的活动数据：一是传统形式的学术活动。它依然扮演着重要角色，如成果发表、获奖、项目申请、会议交流等，在个人学术生涯中，这些形式的学术活动会从表现形式和内容特征两方面产生丰富的活动数据。二是学术相关的社会网络活动。科研的合作性与继承性，使得学者在科研过程中会产生大量的社会网络活动，如与国内外学者的合作，培养的学生学术表现等，由此会产生大量的活动数据。三是表现形式上的活动数据。上述两方面数据不仅存在于传统载体中，网络与通信技术的发展，使这些活动数据以静态或流态的形式大量存在于各种网络平台中，这些数据内容丰富，形式多样。将上述各种内容和形式的数据进行深度聚合，并通过大数据组织、挖掘与关联分析技术，可以全面、深刻而又十分客观地反映学者的学术成果情况。

3.2.2　大数据思维下的学术生涯过程分析　2014 年，Science 的主编 M. McNutt 指出，可以通过科学家学术生涯跟踪来开发更好的学术成果评价指标，如参与跨学科团队的能力，解决复杂问题的创造力和职业道德等[35]。对学术生涯过程分析可以有效判断学者学术积累、学术活跃性、学术发展潜力、学术素养、学术发展导向与前途等，这些要素显然是评价学者科研能力的重要组成部分。甚至通过学术生涯过程分析，可以发现学术研究过程中的成本付出和资源支持等，为学者科研表现的影响因素和本质原因分析提供依据，并进一步为科研管理提供参考。

学术生涯过程的考察分析要根据学者年龄层次进行有针对性地设计“目标导向”，如针对较为年老的学者，主要考察其学术积累、学术贡献等，重在历史维度的学术活动分析；对于年轻学者，主要考察其发展潜力与前途等，淡化对学术积累的考察，以适应哲社学术成果产生的特殊性，重在预见性的学术活动的前瞻分析。此外，学术生涯过程分析要以时间为线索进行“学术存量”与“学术流量”相融合的分析，建立相应的跟踪机制。

大数据思维下学术生涯过程分析，不仅包括对以上各种类型的大量静态数据的整合性分析，还包括对实时动态数据的爬取与处理。其最终的目标导向不仅在于对学者学术研究内在和外在基本状况的明确，更为重要的是发现不同类型学者学术生涯的基本特征，并结合社会发展数据的关联分析，预测其学术研究的发展空间和趋势，实现以学术能力提高为导向的学术评价。

### 3.3　由学术影响力评价转向学术价值和社会效益评价

相比自科学术价值和社会效益的具体性、可描述性，哲社的学术价值更注重意识形态与价值层面的传承和引导，其社会效益更具体验性和个性化，具有一定的模糊性和不确定性。正因如此，哲社学术成果评价更加注重学术共同体的作用。众所周知，学术共同体因学术偏见、认知局限和社会文化等复杂因素的影响，其学术成果评价存在显著的主观性。大数据环境下，通过对理论与应用两方面海量数据的聚合与关联分析，可以客观、深入地发现哲社学术成果的学术价值和社会效益。由此，不仅能够在学术共同体参与评价之前助其完成大量的基础性筛选工作，还能佐证和监督学术共同体的主观倾向，使定性评价更趋于高效和精准。

3.3.1　大数据环境下学术成果学术价值评价的实现

通过云计算、云融合等技术，对跨越不同存储载体的各类型、各结构的科研信息进行整合，并与以往相关成果、当前媒体报道的社会热点和重点相互关联、对比分析，对学术成果的创新性、内容价值进行评价分析；同时，评价分析哲社在跨学科研究、交叉学科研究方法中的应用及其与应用性研究的合作；此外，还可以对学术成果引用或推介者类型以及学术成果引用文献类型等进行分析。通过上述大数据技术和方法来挖掘哲社成果的学术价值，使目前十分流行的以排名列表为表现形式的结果展现转变为对学术价值的深度判断，这种学术价值既可以就其高低分类赋予权重，也可以独具特色，单独评价。

3.3.2　大数据环境下学术成果社会效益评价的实现

20 世纪 50 年代，经济学家开始研究科学研究与试验发展(R&D)对经济增长和生产率的影响[36]，可视为学术成果社会效益评价的雏形；90 年代，社会应用价值被纳入多个领域学术成果评价中，并逐渐引起学者关注[37]；近年来，社会效益评估已成为学术成果评价的

重点，尤其是在健康和医疗领域更为普遍[38]，但是至今没有获得广泛应用，主要原因在于评价所需数据难以获得。目前流行的访谈、案例和文献分析等方法由于时间成本高、数据获取困难等障碍，实际应用价值不高[39]。Altmetrics 的某些指标一定程度上反映了出版物在社会上的影响[40]，但是还不能够深刻并全面地解释学术成果的真实社会作用。因为，在实际应用场景下，关注、点击、下载等指标与成果的实用价值还不能完全等同。此外，高质量研究成果的社会效益也未必高。

大数据为学术成果社会效益评价提供了坚实的数据基础，将大数据思维、技术与方法应用于学术成果的社会效益评价中，会显著增强其可操作性和科学性。大数据时代，通过云搜索、数据集成与整合技术，不仅可以广泛获取存储于各种媒介的全方位科研活动数据，而且还可以将其与社会发展及管理相关的数据相关联，如政府政策性数据、经济发展数据、政治发展数据、公众反馈数据等，考察二者的相关性、分析它们在各种策略、行动指南及政策文献中的被引情况，洞察它们在社会矛盾解决和社会发展趋势预见中的作用，从而有效发现相应学术成果的社会价值。同时大数据思维也有效推动了学术成果评价新指标的开发，如浏览者阅读文献的停留时间、文献被检索频率、搜索指数、文献被二次或多次传播情况等，这些新指标对进一步发现学术成果的社会效益同样重要。

## 4 基于大数据思维的哲社学术成果评价指标体系

评价指标体系是学术成果评价策略的重要依据和落脚点。基于以上分析，我们认为，学术价值和社会效益是哲社学术成果评价的生命线，二者既相互独立又存在必然的联系。学术价值评价的关注点在于“学术”本身，它是为学科体系建设或某一特定研究对象的阐释所提供的支撑性和引导性作用，或者为相关研究提供依据，或者向相关研究输出思想、理论与方法，以此来引导学术的继承与发展；社会效益评价的关注点在于“人与社会”，强调的是学术成果对于社会发展的作用，这种作用可能体现在内化的人的精神或思想文化中，也有可能为社会发展提供具体的指导。学术价值是社会效益的基础，具有良好学术价值的学术成果才有可能具有较高的社会效益，为提升社会效益服务是学术价值的重要指向；社会效益是学术价值向实践转化的重要体现，具有显著社会效益的学术成果通常建立在深厚的学术价值基础之上。

在学术价值中，包括 3 个关键要素：学术创新、学术影响和学术质量。这其中，学术创新完全面向的是学术成果的内容，包括学术成果新颖度（主题、观点与方法）、学术成果前沿性和学术成果前瞻性。新颖度是建立在比较分析基础上的指标，而比较分析需要以分类分析为依据，也即在同一个学科类别内进行比较，如，在同一学科内，该成果的研究内容是否具有突破性（如 h 指数）、跟进性（如 g 指数、e 指数等）或另辟蹊径（如零被引文献研究等）。作为一种创新指标，学术成果前沿性还具有引领特征，即该项成果在主题内容、研究方法和学术观点上能够为某一领域研究起到引导作用，使该领域进一步深化发展（如总体国家安全观中的情报服务研究）。学术成果前瞻性考察的是学术成果是否是对未来预测性的探索，抑或是否是对研究对象本质（潜在性）的挖掘和可发展性的把握。学术成果的新颖度更多地考察某一领域研究对现在的贡献，前沿性和前瞻性则重在对未来的贡献。学术影响一方面是对学术成果传播的表征，另一方面体现了对于其他学术成果或其他人研究行为的影响能力。因此在这一指标中，不仅包含学术迹、会议引用与宣介、学术扩散等表达学术成果传播的指标，也包括学术竞争力、学术话语、学术争鸣等这样的表现学术成果对于其他研究的影响能力的指标。学术质量包括内容质量和形式质量两个部分，前者主要考察学术成果的问题意识、专业性等，后者主要指研究的规范性、论证的完备度等。

学术创新是学术成果的核心，是推动学术成果影响力提升和学术质量持续提升的关键要素；学术影响是学术成果向社会转化的前提，是学术创新的重要目标之一；学术质量是学术创新和学术影响的基础。社会效益包括两个关键要素：社会影响和社会贡献。社会影响强调的是学术成果在社会中的认知、认可，并对社会做出了改变，或为社会创造了经济收益[41]，是社会贡献得以发挥的先导；社会贡献体现的是学术成果对于社会作用发挥的水平和能力。

大数据环境下，要将学者学术活动所产生的所有成果加以关联整合，作为指标量化评估的数据来源，而且不同的指标所应考察的数据源和方法应有所侧重。例如，在学术创新评估中，应将学术成果的内容作为评估对象，这些内容包括主题（选题）、观点、方法等，这些内容创新与否一方面是通过横向比较而得，即对同一学科领域的研究主题（选题）之间的比较、同一研究主题（选题）所提出的观点和采用的方法比较等；另一

方面是通过纵向比较而得，即这些成果是否具有前沿性（处在国际研究的前端）、前瞻性（对潜在问题的洞见与远见）等。学术影响侧重于对成果的传播广度与深度的评估，传播的广度包括传播的速度和传播的空间范围，通过引用、时序分析等可以获得；深度主要指成果对别人相关研究的影响，通过对引用者的学术行为分析可以获得。

基于上述思想，结合指标的互斥性、可操作性、科学性、关联性等原则，本文设计了包含 2 个一级指标、5 个二级指标和 34 个三级指标的哲社学术成果评价指标体系（见表 1）。这一指标体系在实现手段上充分体现了大数据作为信息源（如成果的全文文本、科研政策文本、替代计量指标和重要会议的多类型信息相融合的信息）、大数据作为评价技术（如关联分析、语义挖掘）的应用。此外，这一指标不仅重视全文内容分析（如学术创新中的各类指标等），还关注各种学术活动相关指标（如发表论文、学术会议、媒体发表和 Altmetrics 相关内容等）。

表 1　哲社学术成果评价指标体系

| 一级指标 | 二级指标 | 三级指标 | 观测点 | 实现手段 |
|---|---|---|---|---|
| 学术价值 | 学术创新 | 主题新颖度 | 全文文本 | 语义挖掘[24] |
| | | 观点新颖度 | | 基于内容的引用分析[42－43] |
| | | 方法新颖度 | | 比较分析 |
| | | 前沿吻合度 | | 分类评价[44] |
| | | 成果前瞻性 | | 同行评议 |
| | | 成果交融度 | 成果的学科、主题、参考文献 | 共被引分析、聚类分析、共现分析等[45] |
| | 学术影响 | 学术迹 | 学术成果 | 叶鹰学术迹计算模型[46] |
| | | 会议引用与宣介 | 重要会议的文本、音视频 | 文本挖掘、语义挖掘 |
| | | 学术扩散（时空） | 扩散速度和传播范围 | 被引分析 |
| | | 学术竞争力 | 基础资源、经费投入[47] | 调研分析、问卷分析 |
| | | 学术话语 | 学术观点的传播和引用 | 面向多载体的大数据分析 |
| | | 学术争鸣 | 学术内容被讨论 | |
| | | 传统影响力（总体被引） | 被引、h 指数等 | 被引分析等[48] |
| | | 国际化 | 国际发文及被引 | 文献计量学分析 |
| | | 重要二次文献数据库转载 | 全文转载、文摘转载等 | 《新华文摘》《人大复印资料》等 |
| | 学术质量 | 问题意识 | 与社会发展中问题的衔接 | 文本挖掘、比较分析、同行评议 |
| | | 成果本土化 | 与本国经济社会发展适应性 | |
| | | 重要数据库收录 | 发文量 | SCI、SSCI 等数据库收录 |
| | | 学术获奖 | 获奖成果 | 重要奖项分析 |
| | | 学术规范性 | 全文文本 | 文本结构分析[49]、相关标准规范比较、同行评议 |
| | | 论证完备度 | | |
| | | 持续跟进性 | 学术成果 | 研究内容时空延展分析 |
| | | 专业性[50] | | 同行评议 |
| 社会效益 | 社会影响 | Altmetrics 评分 | Altmetrics. com | Altmetrics 网站获得[51] |
| | | 社会扩散 | 扩散速度与范围 | 各类型和结构的媒体大数据分析 |
| | | 重要媒体刊登/转载 | 媒体显示度 | 《人民日报》《光明日报》等重要媒体分析 |
| | | 与社会热点的相关度 | 社会事件 | 关联分析、同行评议；基于 Altmetrics 的政策文件替代计量分析[52] |
| | | 政策关联/提及/引用 | 政策文本、重大会议数据 | |
| | 社会贡献 | 思想导向 | 主流思想 | |
| | | 成果被采纳 | 被相关机构所采纳 | 与机构相关的多种媒体和形式的大数据挖掘 |
| | | 对自科类成果的支撑 | 自科类成果参考文献 | 基于内容的引用分析 |

在这一指标体系中，我们以哲社成果的特征为基本出发点，强调大数据环境下学术成果呈现载体的多样性，并以量化客观分析方法辅以同行评议方法作为实现手段。但若真正将指标体系落于实践，还需要对各指标重要性进行权重设置，并根据权重值进行标准化、归一化处理，这需要专家调查法和数学、统计学方法的支撑，后续我们将致力于此。此外，鉴于哲社成果学术价值和社会效益呈现的延时性，在利用上述指标

评价时,应采用分阶段、周期性和学术年龄可比较性策略,并且要严格遵循分类评价的原则[53],力求最大化、更客观地描述学术成果状况。同时,大数据环境下,完整的科研行为及科研过程等同样也是学术表现的重要因素,但就目前实际情况而言,这类数据还不能够较为完整地获取,这需要后续相关整合性大数据平台或数据库的支撑,这是后续学术界应该重点探索的问题。

## 5 结语

学术成果评价是推进哲学社会科学健康发展的重要因素,长期以来,传统同行评议和科学计量学方法在哲学社会科学学术成果评价中暴露出的问题越来越多,挑战固有方法的统治地位、开发有针对性的评价策略已刻不容缓。故此,本文提出了大数据环境下哲社学术成果评价的策略以及指标体系,试图在大数据思维指导和大数据方法与技术支撑下,通过全方位科研活动相关数据的互联,以深入内容的大数据相关性分析为核心,促成超越"学术圈内"和传统计量指标的定性与定量方法的有机融合,为大数据在哲社学术成果评价中的应用提供基本策略参考,为哲社学术成果评价在大数据环境下的新指标应用提供依据。

未来,大数据对于哲社学术成果评价的实践应用还有很长的路要走。科研管理部门和学界应在相关大数据平台和数据库建设上多下功夫,如,建立科研活动大数据整合与收集平台,尽可能完备地将学者学术活动的各种数据囊括进来,通过云存储、云服务等方式实现各数据资源的集中管理、分布服务;突破目前各科研数据管理系统、文献数据库之间的界限,实现各系统的互操作,尽可能高效地实现各类学术信息资源共享与智能化搜索;加强全文数据库建设,改变原有仅局限于 PDF 等格式的文献单元处理方式,尽可能规范化、结构化地组织文献中的知识单元,以可以自动化解析的格式处理文献单元等。借此,大数据可以延伸定量分析的范围,监督定性分析的结果,融合补充二者的缺陷和不足,使评价方法更加精准、评价机制趋于完善。



**参考文献:**


[1] 苏新宁. 学术评价与学术考核奖励机制的辩证观[J]. 西南民族大学学报(人文社科版),2017,38(9):1-5.

[2] 李冲,苏永建. 学术评价:量化模式的反思与超越[J]. 自然辩证法研究,2017,33(2):59-63.

[3] 朱剑. 大数据之于学术评价:机遇抑或陷阱? [J]. 中国青年社会科学,2015(4):66-78.

[4] 赵丹群. 定量化学术评价研究的最新进展[J]. 情报理论与实践,2017,40(4):133-137.

[5] 彭秋归. 哲学社会科学评价具有意识形态属性[EB/OL]. [2017-08-12]. http://theory.gmw.cn/2016-04/21/content_19796235.htm.

[6] 姜春林. 人文社会科学研究成果评价的矛盾及其调适[J]. 西南民族大学学报(人文社科版),2015,36(12):233-236.

[7] 代金平. 哲学社会科学发展与学术评价体系[J]. 重庆社会科学,2016(6):13.

[8] 韩毅,夏慧. 时间因素视角下科研人员评价的 Pt 指数研究[J]. 中国图书馆学报,2015,41 (6) : 73-85.

[9] SELVARAJOO K. Measuring merit: take the risk[J]. Science, 2015,347(6218):139-140.

[10] 杨瑞仙,梁艳平. 国内外高校科研评价方法比较研究[J]. 情报杂志,2015,34(9):107-110.

[11] BROOKS T A. Evidence of complex citer motivations[J]. Journal of the American Society for Information Science,1986,37(1):34-36

[12] PINSKI G, NARIN F. Citation influence for journal aggregates of scientific publications: theory with applications to literature of physics[J]. Information processing & management, 1976,12(5): 297-312.

[13] SMALL H. Cited documents as concept symbols[J]. Social studies of science, 1978,8 (3) :327-340.

[14] 丁卫. 引文分析的演变与展望[J]. 情报科学,1992(1):65-72.

[15] 仲伟民. 量化评价扼杀人文学术[J]. 澳门理工学报,2015(3): 100-109.

[16] 王晓莉,叶良均,徐飞,等. SCI 作为科研成果评价标准的局限性研究[J]. 自然辩证法研究,2001 (11):41-47.

[17] GARFIELD E. Can citation indexing be automated? [C]// Proceedings of symposium on statistical association methods for mechanized documentation. Washington: National Bureau of Standards Miscellaneous Publication,1964: 189-192.

[18] 张耀铭. 学术评价存在的问题、成因及其治理[J]. 清华大学学报(哲学社会科学版),2015,30(6):73-88,190-191.

[19] 武夷山,王海燕. 无中生有与唤醒"睡美人"[J]. 情报学报, 2014(2):113.

[20] GLANZEL W, GARFIEL E. The myth of delayed recognition[J]. Scientist,2004,18 (11) :8-9.

[21] VAN RAAANA A F J. Sleeping beauties in science[J]. Scientometrics,2004,59 (3) : 467-472.

[22] 韩毅. 非正式交流回归语境下科技评价的融合路径取向[J]. 中国图书馆学报,2016,42(4):64-74.

[23] TAYLOR M. Exploring the boundaries: how altmetrics can expand our vision of scholarly communication and social impact[J]. Information standards quarterly, 2013,25(2):27-32.

[24] SPAAPEN J, VAN DROOGE L. Introducing 'productive interactions' in social impact assessment[J]. Research evaluation,2011, 20(3):211-218.

[25] SMALL H. Citation context analysis[J]. Progress in communication

science,1982(3):287-312.
[26] 陆伟,孟睿,刘兴帮. 面向引用关系的引文内容标注框架研究[J].中国图书馆学报,2014,40(6):93-104.
[27] ZHANG G, DING Y, MILOJEVIC S. Citation content analysis (CCA): a framework for syntactic and semantic analysis of citation content[J]. Journal of the American Society for Information Science and Technology, 2013,64(7):1490-1503.
[28] SMALL H,KLAVANS R. Identifying scientific breakthroughs by combining co-citation analysis and citation context[C]// Proceedings of the 13th international conference of the International Society for Scientometrics & Informetrics. Leuven: International Society for Scientometrics & Informetrics,2011:783-793.
[29] DING Y. Content-based citation analysis: the next generation in citation analysis[EB/OL].[2017-08-26]. http://www. lis. illinois. edu/events/2012/09 /26/content-based-citation-analysis-next-generation-citation-analysis.
[30] 刘清堂,张思,范桂林,等. 教育大数据视角下的内容语义分析模型及应用研究[J].电化教育研究,2017,38(1):54-61,93.
[31] CRONIN B. Progress in documentation: Invisible colleges and information transfer, are view and commentary with particular reference to the social sciences[J]. Journal of documentation, 1982 (38):212-236.
[32] PRICE J. Little science, big science-and beyond [M]. NewYork: Columbia University Press, 1986.
[33] 黄水清, 王超. 网络环境下非正式信息交流的概念、类别与特点[J]. 图书馆杂志, 2004(6):8-11.
[34] JULIA L. Let's make science metrics more scientific[J]. Nature, 2010, 464 (7288): 488-4899.
[35] MCNUTT M. The measure of research merit[J]. Science, 2014, 346 (6214):1155.
[36] GODIN B,DORE C. Measuring the impacts of science, beyond the economic dimension[EB/OL]. [2017-08-26]. http://www. csiic. ca/PDF/Godin_Dore_Impacts. pdf.
[37] BORNMANN L. What is societal impact of research and how can it be assessed? a literature survey[J]. Journal of the American Society for Information Science and Technology, 2013,64(2):217-233.
[38] HOLBROOK J B. The use of societal impacts considerations in grant proposal peer review: a comparison of five models[J]. Technology & innovation, 2010,12(3): 213-224.
[39] BOAZ A, FITZPATRICK S, SHAW B. Assessing the impact of research on policy: a literature review[J]. Science & public policy, 2009,36 (4):255-270.
[40] BORNMANN L. Alternative metrics in scientometrics: a meta-analysis of research into three altmetrics [J]. Scientometrics, 2015, 103(3):1123-1144.
[41] SANUEL G N,DERRIC G E. Societal impact evaluation: exploring evaluator perceptions of the characterization of impact under the REF2014[J]. Research evaluation,2015,24(3):229-241.
[42] 刘盛博,丁堃,唐德龙.引用内容分析的理论与方法[J].情报理论与实践,2015,38(10):27-32.
[43] 化柏林,王惠临. 语法开发平台的系统设计[J].情报学报, 2006, 25(2):189-196.
[44] 邱均平,赵岩杰,罗力. 科学评价中的论文分类方法研究[J].情报学报,2011(5): 554-560.
[45] 马费成,宋恩梅.我国情报学研究分析:以 ACA 为方法[J]. 情报学报, 2006, 25(3):259-268.
[46] YE F Y,LEYDESFORFF L. The "academic trace" of the performance matrix: a mathematical synthesis of the h-index and the Integrated Impact Indicator ( I3)[J]. Journal of the Association for Information Science and Technology,2014,65(4): 742-750.
[47] 贺德方.我国科技投入效率、效果评价研究[J].情报学报,2006 (6):740-748.
[48] 苏新宁. 图书馆、情报与文献学学术影响力研究报告(2000-2004)——基于 CSSCI 的分析[J].情报学报,2006, 25(2):131-153.
[49] 陆伟,黄永,程齐凯.学术文本的结构功能识别功能框架及基于章节标题的识别[J].情报学报, 2014(9):979-985.
[50] 裴雷,孙建军.中国科技报告质量评价体系与推进策略[J].情报学报,2014(8):813-823.
[51] 李勇,邵钟钰,赵星. Altmetrics 背景下的期刊多维度测评指标研究[J]. 情报学报,2017(2):190-196.
[52] 余厚强,肖婷婷,王曰芬,等.政策文件替代计量指标分布特征研究[J].中国图书馆学报,2017,43(5):57-69.
[53] 武夷山.科研评价指标纵横谈[J].中国科技奖励,2015(3):68-69.

**作者贡献说明:**

李品:负责全文构思和写作;

杨建林:负责修改指导。

## Questions, Strategies and Index Systems on Evaluation of Academic Output for Philosophy and Social Sciences in the Big Data Era

Li Pin Yang Jianlin
School of Information Management, Nanjing University, Nanjing 210023
Jiangsu Key Laboratory of Data Engineering & Knowledge Service, Nanjing University, Nanjing 210023

**Abstract:** [**Purpose/significance**] To solve some defects of peer review and scientometrics methods in the evaluation of philosophy and social sciences academic outputs, the reform and realization of the evaluation of the academic achieve-

ments are explored, especially the evaluation index system of the academic achievements of philosophy and social sciences based on the big data thinking are designed. [**Method/process**] Based on the comparative analysis and comprehensive analysis, this paper analyzes the disadvantages of the traditional philosophy and social sciences evaluation methods, and then analyzes the changes brought by the big data to the philosophy and social sciences evaluation, finally, it puts forward the philosophy and social sciences evaluation strategies and the index systems based on the big data environment. [**Result/conclusion**] In big data era, it is possible to analyze the semantics and its relevance based on all-round academic contents and activity data. By using the citation content and behavior evaluation, the academic activities-centered whole process dynamic evaluation, academic value and social function evaluation, value of academic achievements of philosophy and social sciences can be truly, comprehensively and objectively reflected. Based on above research, an index system for evaluation of academic achievements of philosophy and social sciences is constructed, which is composed of two first-level indicators, five two-level indicators and 34 three-level indicators.

**Keywords**: philosophy and social sciences　academic output　output evaluation　big data



## 2018 年知识管理与知识服务理论、方法与技术学术研讨会在长春召开

2018 年 6 月 26－27 日，由《图书情报工作》杂志社（《图书情报工作》、《知识管理论坛》两刊编辑部）主办，吉林大学管理学院承办的 2018 年知识管理与知识服务理论、方法与技术学术研讨会在吉林大学管理学院隆重召开。

本次会议是一场大咖云集的高端盛会，与会的国内知识管理与知识服务领域重要专家学者与一线工作者就知识管理与知识服务的相关理论、方法与技术进行了探讨，分析和解决当前和未来知识管理与知识服务所面对的新问题与新挑战。

开幕式由《图书情报工作》杂志社社长、主编，中国科学院大学图书情报与档案管理系主任初景利主持。吉林大学管理学院院长赵树宽首先代表承办方致辞，对与会人员表示热烈的欢迎，中国科学院文献情报中心副主任、《图书情报工作》杂志社编委会主任潘亚男致辞，就吉林大学对本次会议的支持表示感谢，并预祝会议圆满召开。大会第一天上午的报告由北京大学图书馆研究馆员、沙特国王图书馆北大分馆馆长朱强主持。台湾资深教授顾敏首先从会场讲台上的“近道、里仁、敦行、恒越”8 个大字讲起，阐释了知识管理的重要性，对知识管理及知识管理系统的概念进行了界定。吉林大学管理学院毕强教授作了“数据驱动发现——知识服务创新的新动能”的报告，他认为数据化将成为图情机构的新业态，数据驱动知识发现将成为知识服务创新的发展方向。南京大学信息管理学院孙建军教授指出，在新技术和科研新范式产生的复杂环境的变化下，图书情报工作应该重塑战略思维，以知识服务为重点并结合其他方面的发展，来培育图书情报领域新动能。行者互联创始人，创新与知识管理联盟秘书长吴庆海作了知识萃取技术与案例实务的报告，从知识管理实践的三次大迭代谈起，探讨了知识萃取的意义和相关技术与方法，并以浅显易懂的案例生动形象地说明了知识萃取在生活、学习和工作等各个方面的重要性和使用技巧。

会议第一天下午的报告由吉林大学管理学院教授、博士生导师黄微主持。南京大学信息管理学院教授苏新宁、武汉大学信息管理学院教授陆伟、蓝凌软件股份有限公司副总裁与研究院院长夏敬华、北京大学光华管理学院教授董小英、郑州航空工业管理学院信息科学学院院长刘永教授、南开大学商学院副院长兼信息管理系主任李月琳教授分别作了知识管理相关主题报告。其中苏新宁教授从知识组织的角度、陆伟教授从语义标注和知识挖掘的角度、刘永教授从知识管理与知识服务主要驱动技术的角度对知识管理的实现技术等进行了交流。董小英教授对大数据时代隐性知识的价值及其显性化实现进行了汇报。李月琳教授从信息检索系统的评估延伸到对知识服务平台评估的研究，对评估的内容、标准及其应用进行了分析。

会议第二天上午由东北师范大学图书馆馆长刘万国主持报告。南开大学商学院信息资源管理系教授柯平将知识服务分为前知识服务时代和后知识服务时代，并对这两个时代的特点进行了分析，提出了知识服务未来发展的三个方向。南京大学图书馆副馆长、保密学院教授邵波回顾了南京大学图书馆系统的建设历程，以新一代系统中央知识库的构建与应用为核心议题，介绍了目前南京大学相关的实验项目，并详细报告了中央知识库的构建框架和数据处理流程及相关标准等内容。南京农业大学社科处处长黄水清则详细介绍了其团队在《春秋经传》人物识别及关联分析项目中的研究成果。华中师范大学信息管理学院院长李玉海从图书馆智慧服务释义、智慧服务中的数据恋爱模型及数据的获取与刺激三个方面，对图书馆智慧服务中的数据获取进行了报告。重庆大学图书馆馆长杨新涯作了题为“不忘初心：图书馆目录的发展创新”的主题报告，对数字文献资源目录的来源、文献分类、标引与著录和 OPAC＋四个关键问题进行了阐述，并介绍了重庆大学在数字资源的组织和相关应用方面的实践。

最后，吉林大学管理学院副院长、教授王晰巍主持闭幕式，《图书情报工作》杂志社社长、主编，中国科学院大学图书情报与档案管理系主任初景利做了本次会议的闭幕总结。